PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :		(11) International Publication Number:	WO 96/17942
C12N 15/62, 9/88, C07K 14/60, 14/605, 14/635, 1/30, 1/113	A1	(43) International Publication Date:	13 June 1996 (13.06.96)

(21) International Application Number: PCT/US95/15800

(22) International Filing Date: 7 December 1995 (07.12.95)

(30) Priority Data:
08/350,530 7 December 1994 (07.12.94) US

(71) Applicant: BIONEBRASKA, INC. [US/US]; 3820 Northwest 46th Street, Lincoln, NE 68504 (US).

(72) Inventors: PARTRIDGE, Bruce, E.; 1209 South 25th Street, Lincoln, NE 68502 (US). STOUT, Jay, S.; 1921 Sewell, Lincoln, NE 68502 (US). HENRIKSEN, Dennis, B.; Apartment 723, 343 North 44th Street, Lincoln, NE 68503 (US). MANNING, Shane, D.; Apartment 2217, 11206 Seward Plaza, Omaha, NE 68154 (US). DE LA MOTTE, Rebecca, S.; 1209 South 25th Street, Lincoln, NE 68502 (US). HOLMQUIST, Barton; 442 West Lakeshore Drive, Lincoln, NE 68526 (US). WAGNER, Fred, W.; Route 1, Box 77B, Walton, NE 68461 (US).

(74) Agent: BRUESS, Steven, C.; Merchant, Gould, Smith, Edell, Welter & Schmidt, 3100 Norwest Center, 90 South Seventh Street, Minneapolis, MN 55402 (US).

(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, LS, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PRODUCTION OF PEPTIDES USING RECOMBINANT FUSION PROTEIN CONSTRUCTS

(57) Abstract

A method for the isolation and/or purification of a recombinant peptide by employing a fusion protein construct which includes a carbonic anhydrase and a variable fused polypeptide is provided. The method includes precipitating either the fusion protein construct or a fragment of the fusion protein construct including the carbonic anhydrase. Inclusion bodies which includes the fusion protein construct and a method of producing a peptide which includes expressing the fusion protein construct as a part of an inclusion body are also provided. Fusion protein constructs which include a carbonic anhydrase and certain target peptides are also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

PRODUCTION OF PEPTIDES USING RECOMBINANT FUSION PROTEIN CONSTRUCTS

Background of the Invention

In vitro DNA manipulation and the attendant transfer of genetic information have developed into a technology that allows the efficient expression of endogenous and foreign proteins in microbial hosts. Recombinant DNA techniques have made possible the selection, amplification and manipulation of expression of the proteins and peptides.

Although expression of any foreign protein in any microbial host is theoretically possible, the stability of the protein produced often limits such practice and results in a low yield. In particular, small foreign proteins and oligopeptides are not easily overproduced in most cellular hosts. Expression of a small peptide in a host cell raises the possibility that the host will assimilate the polypeptide.

In response to this problem, small peptides have been expressed as fusion proteins with a second larger peptide, such as a peptide marker (e.g., beta-galactosidase or chloramphenicol acetyl transferase). While the resulting fusion protein may not be readily assimilated, isolation and purification of the desired protein is often not very efficient or effective.

The development of methods which permit the isolation and/or purification of recombinant peptides without requiring the use of any column chromatography separations, e.g., chromatography on an affinity column, would be particularly desireable. Such columns are typically expensive and may have a relatively limited lifetime. Accordingly, there is a continuing need for flexible, convenient methods for the isolation and purification of recombinant peptides.

Summary of the Invention

The present invention provides a method for the production of a recombinant peptide which employs a fusion protein construct including a carbonic anhydrase

and a variable fused polypeptide. Typically, the carbonic anhydrase is a mammalian carbonic anhydrase (mCA). The carbonic anhydrase and the variable fused polypeptide may be linked together by a cleavage site. The cleavage site is an amino acid or sequence of amino acids which results in the protein construct being selectively cleaved on treatment by a cleavage agent. The cleavage agent may be a chemical reagent, e.g. cyanogen bromide, which recognizes a certain chemical cleavage site such as a methionine residue. The cleavage agent may be also be an endopeptidase which cleaves the construct at a specific point in relation to a particular amino acid or sequence of amino acids ("an enzymatic cleavage site").

One embodiment of the invention is directed to a method of producing a peptide which includes precipitating the fusion protein construct. The precipitated construct may be resolubilized. After cleavage, the fragment containing the carbonic anhydrase ("carbonic anhydrase fragment") may be removed by chromatography, filtration or, preferably, by precipitation.

A second embodiment provides a method of producing a peptide which includes (i) cleaving the fusion protein construct to produce a soluble carbonic anhydrase fragment and a soluble variable fused polypeptide fragment, and (ii) precipitating the carbonic anhydrase fragment. The carbonic anhydrase fragment includes the carbonic anhydrase and may also include other amino acids residues, e.g., from an interconnecting peptide or a chemical moiety created by the cleavage reaction. The variable fused polypeptide fragment includes at least one copy of a target peptide and may also include additional amino acid residues such as an N-terminal or C-terminal tail sequence or an intraconnecting peptide. Where the variable fused polypeptide fragment includes more than one copy of a target peptide, the target

peptides are typically linked by an intraconnecting peptide which includes a cleavage site.

A third embodiment of the invention provides a method of producing a peptide which includes expressing the fusion protein construct as part of an inclusion body in a host cell, e.g., an *E. coli* host cell, and isolating the fusion protein construct. The fusion protein construct typically includes at least 2 copies of a target peptide. However, if the fusion protein construct includes target peptide corresponding to a peptide selected from the group consisting of GRF(1-44) (SEQ ID NO:20), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22), a construct having only a single copy of the target peptide may be expressed as a part of an inclusion body. GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22) represent peptides having the designated amino acids corresponding to growth hormone releasing factor, glucagon-like peptide 1 and parathyroid hormone respectively. The sequences for amino acids 1-44 of growth hormone releasing factor (GRF(1-44) (SEQ ID NO:20)) and amino acids 1-36 of Glucagon-like Peptide 1 (GLP1(1-36) (SEQ ID NO:24)) are disclosed in International Application No. PCT/US94/08125, the disclosure of which is incorporated herein by reference. The sequence for amino acids 1-84 of human parathyroid hormone (PTH(1-84) (SEQ ID NO:25)) is disclosed in Hendy et al., Proc.Natl.Acad.Sci.,USA, 78, 7365 (1981) and T. Kimura et al., BBRC, 114, 493 (1983), the disclosure of which is incorporated herein by reference.

Another embodiment of the invention provides a method of producing a peptide which includes cleaving the fusion protein construct to produce a carbonic anhydrase fragment and a variable fused polypeptide fragment. The carbonic anhydrase fragment and the variable fused polypeptide fragment are precipitated. The precipitated fragments may be extacted with with a

solvent which includes an organic component to recover the variable fused polypeptide fragment.

The invention also provides an inclusion body containing the fusion protein construct expressed in a host cell, e.g., an *E. coli* host cell. The fusion protein construct includes a carbonic anhydrase and a target peptide. The target peptide preferably includes at least one copy of an amino acid sequence corresponding to a peptide selected from the group consisting of GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22).

Finally, the invention also provides a fusion protein construct which includes the carbonic anhydrase and a target peptide. The target peptide includes an amino acid sequence corresponding to a peptide selected from the group consisting of GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22).

Brief Description of the Figures

FIG. 1 depicts the various formulas for variable fused polypeptides formed of multiple units of target peptides.

Detailed Description of the Invention

The expression of a foreign protein by interaction of recombinant vectors with the biosynthetic machinery of host cells or host organisms is a well-known technique for biochemical protein synthesis. The present invention utilizes a novel modification of this expression technique in combination with a carbonic anhydrase-based fusion protein construct to establish a new, low cost, highly efficient method for the large scale biological synthesis of peptides. The method makes it possible to design processes for the isolation and/or purification of a recombinant peptide based solely on chemical methods, i.e. without requiring affinity chromatography or other chromatographic steps.

The fusion protein construct also permits the use of a ligand immobilized affinity separation technique if desired. The method eliminates expensive machinery and reagents, long synthetic times, low reaction efficiency and produces fewer faulty copies and a higher yield of the product peptide, than the solid phase peptide synthesis.

The method of the invention incorporates a combination of cellular factors and biologic compositions that enable ready expression of foreign proteins by biological systems. Because the expressed fusion protein construct incorporates a binding protein, carbonic anhydrase, subversion of the particular identity of the target peptide by the expression mechanism of the cell or organism is not permitted. Variation in expression efficiency is minimized. The preferred inductive expression mechanism allows production of peptide constructs that might otherwise be toxic to host cells.

The method of the invention also incorporates factors that contribute substantially to the efficiency, capacity and yield of the purification technique. The solubility properties of the carbonic anhydrase and fusion construct including the carbonic anhydrase facilitate a clean, complete separation of the variable fused polypeptide from other constituents. The relatively low molecular weight of the carbonic anhydrase permits a high capacity and efficiency per unit weight of the fusion construct. Because the method does not require any special reagents or reactions, the production of undesirable side products is avoided.

The size of the fusion protein construct will vary depending on the nature and number of copies of the target peptide. The fusion protein construct is large enough to avoid degradation by the host cell (e.g., at least about 60 to 80 amino acid residues) and not so large that it can not be effectively expressed by the

host cell. As a practical matter, the fusion protein construct will have a molecular weight of up to about 500,000 although larger constructs are also within the scope of the present invention. The size of the fusion protein construct is chosen such that it may be expressed by the host cell so as to avoid introducing errors in the protein sequence. This places practical limitations on the number of copies of the target peptide present in a given construct. The actual number will vary depending on the size and nature of a particular target peptide within the limitations set by the factors discussed above.

Carbonic Anhydrase

The carbonic anhydrase may be derived from a variety of sources. Suitable sources include vertebrates and in particular mammalian sources, e.g., a human carbonic anhydrase (hCA), a rat carbonic anhydrase, a feline carbonic anhydrase, an equine carbonic anhydrase or a bovine carbonic anhydrase. A carbonic anhydrase derived from other mammalian species such as ovine, murine, porcine and monkey, may also be used to form the fusion protein construct of the present invention. An example of a suitable carbonic anhydrase is human carbonic anhydrase II ("hCAII"; see, Taylor et al., Biochemistry, 9, 2638 (1970)).

The carbonic anhydrase may also be a modified functional version of a carbonic anhydrase. The modified functional version will retain a carbonic anhydrase's capability of being precipitated as a function of solution conditions, e.g. at a particular pH or as a function of metal ion or salt concentration. Modified versions which can be precipitated from a wide variety of solutions by adjusting the pH of the solution to between about 3.2 and about 6.0, more preferably to between about 3.5 and about 5.5, are preferred. The modified functional version typically also retains the

ability of a carbonic anhydrase to strongly bind with the inhibitors, benzenesulfonamide, acetazolamide, or derivatives thereof. A detailed discussion of the binding properties of carbonic anhydrases with respect to sulfanilamide and acetazolamide inhibitors is set forth in International Application No. PCT/US91/04511, which is hereby incorporated by reference.

Examples of suitably modified carbonic anhydrases include functional substitution mutants which (I) do not contain methionine, (II) have all or some glutamates replaced by another negatively charged amino acid, preferably aspartate, (III) have all or some arginines replaced by another positively charged amino acid, preferably lysine, (IV) have all or some asparagines replaced by another amino acid, preferably glutamine, (V) have methionine replaced by another amino acid, preferably alanine, serine, cysteine, threonine or leucine, or (VI) have cysteine replaced by another amino acid, preferably serine, threonine, leucine or alanine.

Examples of suitably modified versions of carbonic anhydrase also include polypeptides representing a functional fragment of a carbonic anhydrase such as hCAII, such as but not limited to hCAII C-terminated at cysteine 205, asparagine 231, methionine 240, or leucine 250 or N-terminated at proline 21 or glycine 25. The functional fragment retains the ability of a carbonic anhydrase to be precipitated from a wide range of solutions by adjusting the pH.

Other suitably modified versions include substitution mutants of hCAII and substitution mutants of the functional fragments of hCAII having the following substitutions:

i. hCAII with all or some glutamate amino acid residues replaced by another negatively charged amino acid (AA), preferably an aspartic acid residue.

ii. hCAII with all or some arginine AA residues

replaced by another positively charged AA, preferably a lysine residue.

iii. hCAII or a functional fragment thereof with one or more of the following modifications to the AA positions at [N11X, G12X] (asparagine glycine), [N62X, G63X], [N231X, G232X], M240X (methionine), or C205X (cysteine) modified as follows (where those AA positions are present in a given fragment): the asparagine is changed to glutamine or glycine is changed to alanine, methionine is changed to alanine, serine, cysteine, threonine or leucine, and cysteine is changed to serine threonine, leucine or alanine.

The carbonic anhydrase and fusion proteins including the carbonic anhydrase preferably are capable of being precipitated from a wide range of solutions by adjusting the pH of the solution to between about 3.2 and about 6.0, preferably between about 3.5 and about 5.5, and more preferably between about 4.0 and about 5.0. Examples of solutions from which the carbonic anhydrase can be precipitated using this method include solutions containing polyethyleneimine, urea (e.g., at least about 2M) or guanidine hydrochloride (e.g., at least about 2M). Preferably, the carbonic anhydrase remains soluble in a wide variety of solutions outside of the pH range of about 3.2 to about 6.0.

Variable Fused Peptide

The variable fused polypeptide may have several forms (see e.g., Figure 1). One includes a single target peptide. A second is composed of multiple tandem units of a single target peptide that are linked by an amino acid or an intraconnecting peptide. In this case, the intraconnecting peptide usually but not necessarily differs in structure and selectivity from the interconnecting peptide (which links the carbonic

anhydrase and the variable fused polypeptide). When different, the intraconnecting peptide nevertheless has the same general function as the interconnecting peptide so that two different cleavage agents of an enzymatic or chemical nature will separately cleave the variable fused polypeptide from the carbonic anhydrase and cleave the individual target peptides from each other. The third form is a single unit composed of several (i.e., two or more) identical or different target peptides tandemly interlinked together by innerconnecting peptides. The fourth form is composed of repeating multiple tandem units linked together by intraconnecting peptides wherein each unit contains the same series of different individual target peptides joined together by innerconnecting peptides. The fifth form is composed of a series of tandem units linked together by intraconnecting peptides wherein each unit contains several identical or different target peptides joined by innerconnecting peptides and the target peptides do not repeat from unit to unit. The sixth form is composed of identical multiple tandem units wherein each unit contains several identical target peptides joined by innerconnecting peptides.

The target peptide may incorporate all or a portion of any natural or synthetic peptide desired as a product, e.g., any desired protein, oligopeptide or small molecular weight peptide. For the purposes of this application a peptide includes at least two amino acid residues linked by a peptide bond. Suitable embodiments of the target peptide which may appear as single or multiple linked units in the variable fused polypeptide include caltrin, calcitonin, insulin, tissue plasminogen activator, growth hormone, growth factors, growth hormone releasing factors, erythropoietin, interferons, interleukins, oxytocin, vasopressin, ACTH, collagen binding protein, parathyroid hormone, glucagon like peptide, glucagon, proinsulin, tumor necrosis

factor, substance P, brain naturetic peptide, individual heavy and light antibody chains, individual antibody chain fragments especially such as the isolated variable regions (VH or VL) as characterized by Lerner, Science, 246, 1275 et. seq. (Dec. 1989) and epitopal regions such as those characterized by E. Ward et al., Nature, 341, 544-546 (1986) wherein the antibodies, chains, fragments and regions have natural or immunogenetically developed antigenicity toward antigenic substances. Additional embodiments of the desired polypeptide include polypeptides having physiologic properties, such as sweetening peptides, mood altering polypeptides, nerve growth factors, regulatory proteins, functional hormones, enzymes, DNA polymerases, DNA modification enzymes, structural polypeptides, neuropeptides, polypeptides exhibiting effects upon the cardiovascular, respiratory, excretory, lymphatic, immune, blood, reproductive, cell stimulatory and physiologic functional systems, leukemia inhibitor factors, antibiotic and bacteriostatic peptides (such as cecropins, attacins, apidaecins), insecticidal, herbicidal and fungicidal peptides as well as lysozymes.

Among suitable variable fused polypeptides are peptides which include an amino acid sequence having the formula:

$$-\text{TargP-(CS2)-[-(Ln1)}_n\text{-(CS1)}_m\text{-TargP-(CS2)-]}_r$$

wherein the -CS1- and -CS2- are cleavage sites, the -(Ln1)- is a linking sequence, the -TargP- is the target peptide, n and m are 0 or 1, and r is an integer from 1 to about 150. The target peptide may include an amino acid sequence corresponding to GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) or PTH(1-34) (SEQ ID NO:22). The -(CS1)- and -(CS2)- cleavage sites may be a chemical cleavage site or an enzymatic cleavage site. The enzymatic cleavage site may be recognized by an endopeptidase or by an exopeptidase (e.g., where r is 1 and the -(CS2)- is the C-terminal residue of the

construct). Preferably, the enzymatic cleavage site is recognized by an endopeptidase.

Cleavage Site

The fusion protein construct may include a chemical cleavage site or an enzymatic cleavage site. The cleavage site or sites which may be incorporated into the fusion protein construct will depend upon the identity of the target peptide(s) present. The cleavage site and target peptide are typically selected so that target peptide does not contain an amino acid sequence corresponding to the cleavage site. Secondary considerations will also influence the choice of a particular cleavage site. In some instances, the cleavage sites may be designed so as to avoid the use of a enzymatic cleavage reaction. This may be accomplished by employing a chemical cleavage site, such as a site which may by cleaved after treament with an S-cyanylating agent (e.g, 2-nitro-5-thiocyanatobenzoate) or by treatment with an acid having a pK_a of no more than about 3.0. In other instances, it may be desireable to employ a cleavage site which permits a modification of the target peptide to introduce a specific functional group, e.g., a C-terminal α-carboxamide group.

Chemical and enzymatic cleavage sites and the corresponding agents used to effect cleavage of a peptide bond close to one of these sites are described in detail in International Application Nos. PCT/US91/04511 and PCT/US94/08125, the disclosure of which is herein incorporated by reference. Examples of peptide sequences (and DNA gene sequences coding therefor) suitable for use as cleavage sites in the present invention and their corresponding cleavage enzymes or chemical cleavage conditions are shown in Table 1 below. The gene sequence indicated is one possibility coding for the corresponding peptide sequence. Other DNA sequences may be constructed to code for the same peptide sequence.

Table 1

Enzymes for Cleavage	Peptide Sequence	DNA Sequence
Enterokinase	$(Asp)_4Lys$ (SEQ ID NO:2)	GACGACGACGATAAA (SEQ ID NO:1)
Factor Xa	IleGluGlyArg (SEQ ID NO:4)	ATTGAAGGAAGA (SEQ ID NO:3)
Thrombin	GlyProArg or GlyAlaArg	GGACCAAGA or GGAGCGAGA
Ubiquitin Cleaving Enzyme	ArgGlyGly	AGAGGAGGA
Renin	HisProPheHisLeu-LeuValTyr (SEQ ID NO:6)	CATCCTTTTCATC-TGCTGGTTTAT (SEQ ID NO:5)
Trypsin	Lys or Arg	AAA OR CGT
Chymotrypsin	Phe or Tyr or Trp	TTT or TAT or TGG
Clostripain	Arg	CGT
S. aureus V8	Glu	GAA

Chemical Cleavage	Peptide Sequence	DNA Segence
(at pH3)	AspGly or AspPro	GATGGA or GATCCA
(Hydroxylamine)	AsnGly	AATCCA
(CNBr)	Methionine	ATG
BNPS-skatole	Trp	TGG
2-Nitro-5-thiocyanatobenzoate	Cys	TGT

Production of Recombinant Peptides

Host cells transformed with an expression vector carrying a recombinant fusion protein construct gene may be employed to express the fusion protein construct. The host cell may be a prokaryotic or eukaryotic host cell or a cell in a higher organism. In one preferred embodiment of the invention, the host cell is a microbial host cell such as *E. coli*. The vector carrying the protein purification construct gene may be prepared by insertion of the DNA segments coding for the fusion protein construct into an appropriate base vector. Methods for expression of single- and multicopy recombinant fusion protein products are disclosed in International Application No. PCT/US91/04511, the disclosure of which is incorporated herein by reference.

The expression vector incorporates the recombinant gene and base vector segments such as the appropriate regulatory DNA sequences for transcription, translation, phenotyping, temporal or other control of expression, RNA binding and post-expression manipulation of the expressed product. Structural features such as a promoter, an operator, a regulatory sequence and a transcription termination signal are generally included in the expression vector. The expression vector may be synthesized from any base vector that is compatible with the host cell or higher organism and provides the foregoing features. The regulatory sequences of the expression vector will be specifically compatible or adapted in some fashion to be compatible with prokaryotic or eukaryotic host cells or higher organisms. Post-expression regulatory sequences, which cause secretion of the polypeptide construct can be included in the eukaryotic expression vector. It is especially preferred that the expression vector exhibit a stimulatory effect upon the host cell or higher organism such that the polypeptide construct is overproduced relative to the usual biosynthetic expression of the host.

Typically, the recombinant gene is inserted into an appropriate base vector which is used to transform host cells or higher organisms with the resulting recombinant vector. The fusion protein construct may be expressed within the host cell or higher organism as a soluble product or as a product that is insoluble in the cell. Alternatively, the fusion protein construct may be expressed as a secreted product by the host cell or higher organism. Preferably, the fusion protein construct may be expressed in a host cell such as *E. coli* as a part of an inclusion body, i.e., an aggregate of insoluble material. Fusion protein constructs which are expressed as part of an inclusion body typically include two or more copies of a target peptide. It has been found that most single copy fusion protein constructs based on a carbonic anhydrase are expressed as a soluble product. For example, single copy fusion protein constructs which include hCAII and a single copy of a target peptide, such as calcitonin, substance P, angiotension, ATPase-IP, ɛSU ATPase, and asparagine synthetase, are expressed as soluble cell products in *E. coli*. Surprisingly, it has been discovered that in some instances, however, a single copy carbonic anhydrase-based fusion protein may be expressed as part of an inclusion body. In particular, fusion proteins which include hCAII and a single copy of GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) or PTH(1-34) (SEQ ID NO:22) are expressed as a part of an inclusion body in an *E. coli* host cell such as *E. coli* BL21 F^{-} ompT r_B^{-} m_B^{-} (DE3).

The present invention provides a method of producing a recombinant peptide which includes precipitating either the fusion protein construct or a fragment of the fusion protein construct including the carbonic anhydrase (collectively a "CA-based protein"). hCAII is known to precipitate when a solution at about neutral pH of the enzyme in its native form is

acidified. It is believed that lowering the pH leads to a zinc cation being stripped out of the native enzyme and substantial conformational changes resulting in its precipitation. Surprisingly, it has been found that carbonic anhydrase and, in particular, a mammalian carbonic anhydrase such as hCAII, exhibits this same solubility behavior as a function of pH regardless of whether the enzyme is in its native form. Moreover, it has been discovered that fusion proteins incorporating a carbonic anhydrase may be precipitated in a similar fashion as a function of pH. A CA-based protein typically may be precipitated from a solution simply by adjusting the pH to between about 3.2 and about 6.0, preferably between about 3.5 and about 5.5, and more preferably between about 4.0 and about 5.0. CA-based proteins may also be precipitated from solutions having a high salt concentration, e.g., from solutions which include at least about 5% (w/v) sodium sulfate or at least about 5% (w/v) ammonium sulfate.

In one embodiment of the invention, a recombinant peptide is isolated by a process which includes precipitating the fusion protein construct. The precipitation step may be carried out with a fusion protein construct which has been expressed as a soluble product or with a fusion protein construct which has been expressed as a part of an inclusion body. In the latter instance, the inclusion body is dissolved prior to the precipitation step, e.g., in a solution which includes at least about 0.5M citric acid. The fusion protein construct may also be dissolved in a variety of other solutions, such as a solution containing a chaotropic agent (e.g., guanidine hydrochloride) or a detergent (e.g., SDS), a solution containing at least about 2 M urea or a solution having a pH of at least about 10.0.

The fusion protein construct may be precipitated from the above solutions by adjusting the pH of the

solution to between about 3.2 and about 6.0 and preferably to between about 3.5 and about 5.5. The efficiency of the precipitation may be enhanced in some cases by adding a salt to the solution in addition to adjusting the pH. Suitable salts include sodium chloride, potassium chloride, manganese sulfate, sodium sulfate, sodium acetate, and lanthanum chloride. The addition of divalent metal ions, such as Mg^{2+}, Ca^{2+}, Zn^{2+}, Sr^{2+} or Co^{2+} together with the pH adjustment may also enhance the efficiency of the precipitation. If the solubilized fusion protein construct is precipitated from an alkaline solution, the pH of the solution is typically adjusted by adding an acid such as acetic acid, formic acid, citric acid, phosphoric acid, and hydrochloric acid. As an alternative to precipitation based on pH adjustment, the solubilized fusion protein construct may be also precipitated from solution simply by the addition of a sufficient amount of a salt such as sodium sulfate or ammonium sulfate.

In some instances, it may be useful to remove cell debris and nucleic acid material from a crude cell lysate as an initial step in the isolation of a recombinant peptide. This is typically accomplished by adding at least about 0.25% and preferably at least about 0.38% polyethyleneimine (PEI) to the lysate. The addition of the PEI generally causes cell debris and about 90-95% of the DNA present to be precipitated. The mCA-based fusion protein may then be precipitated from the supernatant simply by adjusting the pH of the PEI solution to between about 3.5 and about 6.0.

The precipitated fusion protein may be resolubilized, e.g., by dissolution in a solution having a pH of at least about 10, and preferably at least about 10.5. The precipitated fusion protein may also be redissolved in other types of solutions, such as solutions which include a chaotropic agent, a detergent or at least about 0.5 M citric acid. After cleavage,

the fragment containing the carbonic anhydrase ("carbonic anhydrase fragment") may be removed by chromatography, filtration, or preferably, by precipitation. As noted above, the carbonic anhydrase fragment may be precipitated simply by again adjusting the pH of the solution to between about 3.5 and about 6.0. If desired, as an alternative, the carbonic anhydrase fragment may be separated from the cleavage products using a ligand immobilized affinity separation technique. For example, the carbonic anhydrase fragment, in native or renatured form, may be removed by contacting a solution of the cleavage products with an inhibitor (e.g., an affinity column) which includes benzenesulfonamide compound or an acetazolamide compound. Preferably the benzenesulfonamide compound includes a sulfanilamide compound such as an amino substituted benzenesulfonamide (e.g., 4-amino-benzenesulfonamide) or a derivative thereof. The acetazolamide compound is 5-acetamido-1,3,4-thiadiazol-2-sulfonamide or a derivative thereof.

In another embodiment of the invention, the method of producing a recombinant peptide includes (i) cleaving the fusion protein construct to produce a soluble carbonic anhydrase fragment and a soluble variable fused polypeptide fragment, and (ii) precipitating the carbonic anhydrase fragment. The carbonic anhydrase fragment may include other amino acids residues in addition to the carbonic anhydrase. These other amino acid residues typically are derived from an interconnecting peptide present in the fusion protein construct. Suitable interconnecting peptides for use in the present invention include amino acid sequences containing a chemical or enzymatic cleavage site. After cleavage, the carbonic anhydrase fragment typically includes at least a portion of the interconnecting peptide or a derivative thereof. The derivative generally is produced as a byproduct of the cleavage

reaction. For example, a carbonic anhydrase fragment having a C-terminal homoserine residue may be created by cleavage of a fusion protein construct at the C-terminal peptide bond of a methionine residue.

The supernatant fraction remaining after the precipitation of the carbonic anhydrase fragment may be isolated by centrifugation. The supernatant fraction contains the variable fused polypeptide fragment and, if desired, the variable fused polypeptide fragment purified further by conventional methods used for the isolation of peptides.

To enhance the recovery of the variable fused polypeptide fragment, the precipitated carbonic anhydrase fragment may be extracted with a solvent which includes an organic solvent. Suitable solvents include an organic component such as acetonitrile, propanol, citric acid, polyethyleneglycol, or mixtures thereof. For example, aqueous solutions including an organic solvent, e.g., 50% aqueous acetonitrile and 50% aqueous n-propanol, may be used to carry out the extraction.

Prior to being subjected to the cleavage reaction, the fusion protein construct may be purified by a variety of methods. For example, the fusion protein construct may be purified using a ligand immobilized affinity separation technique or by precipitation. Purification of the fusion protein construct by precipitation is preferred as it allows the recombinant fusion protein construct to be produced by a process which relies solely on wet chemical separation processes such as precipitation, filtration or centrifugation without the need for renaturation or affinity purification.

A third embodiment of the invention provides a method of producing a recombinant peptide which includes expressing the fusion protein construct as a part of an inclusion body in an *E. coli* host cell, and isolating the fusion protein construct. The fusion protein

construct typically includes multiple copies of a target peptide. As noted above however, it is possible to express fusion protein constructs which include an mCA and a single copy of an amino acid sequence corresponding to GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22). Examples of suitable single copy fusion proteins which may be expressed as inclusion bodies in an *E. coli* host cell are inclusion fusion proteins which include hCAII and a single copy of any one of GRF(1-41) (SEQ ID NO:23), GRF(1-44) (SEQ ID NO:20), GLP1(1-34) (SEQ ID NO:26), GLP1(7-34) (SEQ ID NO:21), GLP1(7-36) (SEQ ID NO:27), GLP1(7-37) (SEQ ID NO:28), PTH(1-34) (SEQ ID NO:22), PTH(1-38) (SEQ ID NO:29) and PTH(1-84) (SEQ ID NO:25).

The inclusion bodies may be isolated from a crude cell lysate by conventional techniques, e.g., by centrifugation. The crude inclusion bodies may be subjected to an initial purification step such as washing the inclusion bodies with a 50mM Tris, 1mM EDTA, pH 7.8 solution or a 100mM EDTA solution to remove any cell debris and/or nucleic acid materials present.

The inclusion bodies may be dissolved under a variety of conditions. For example, the inclusion bodies may be dissolved under acidic conditions in a solution having a pH of no more than about 3.2 (e.g., in a solution which includes at least about 500mM citric acid). Alternatively, the inclusion bodies may be dissolved in a solution which has a pH of at least about 10. The dissolution of the inclusion bodies under basic conditions may be facilitated by the addition of a small amount of a surfactant, such as N-lauroyl sarcosine, to the solution. The inclusion bodies may also be dissolved in a solution which includes an effective amount of a chaotropic agent such as 6N HCl or urea, or a detergent such as N-lauroyl sarcosine, sodium dodecyl sulfate, sodium octyl sulfate or cetyl trimethyl ammonium bromide (CTAB). It has unexpectedly been

discovered that after the inclusion bodies have been dissolved in the solutions described above, the solubilized fusion protein construct may be precipitated by adjusting the pH of the solution to between about 3.5 and about 6.0, and preferably to between about 4.0 and about 5.0. As an alternative, the solubilized fusion protein construct may be isolated by a process which includes a ligand immobilized affinity separation technique.

Another embodiment of the invention includes cleaving the fusion protein construct to produce a carbonic anhydrase fragment and a variable fused polypeptide fragment and precipitating both the carbonic anhydrase fragment and the variable fused polypeptide fragment. The precipitated fragments may be extacted with a solvent which includes an organic component to recover the variable fused polypeptide fragment. Suitable solvents include an organic component such as acetonitrile, propanol, citric acid, polyethyleneglycol, or mixtures thereof.

The invention will be further described by reference to the following detailed examples.

Example 1. Description of the Expression System

Origin, phenotype and genotype of the host cells.

The bacterial host for expression, *E. coli* BL21 F$^-$ ompT $r_B^- m_B^-$ (DE3) was obtained from Novagen, Inc., Madison, WI. These *E. coli* cells gave high levels of expression of genes cloned into expression vectors containing the bacteriophage T7 promoter. Bacteriophage (DE3) which contains the T7 RNA polymerase gene has been integrated into the chromosomal DNA of the BL21 (DE3) cells. The T7 RNA polymerase gene is controlled by the lacUV5 promoter and the lacI gene product.

Construction of pBN1

An expression vector, pET31F1mhCAII containing the

hCAII gene was obtained from Dr. P.J. Laipis at the University of Florida. The pET31F1mhCAII was prepared as described by Tanhauser et al., Gene, 117, 113 (1992). Plasmid pET31F1mhCAII contains the coding region for hCAII (human carbonic anhydrase II) downstream of a bacteriophage T7 promoter in a pUC-derived plasmid backbone. Two synthetic oligonucleotides, 5'-A GCT TTC GTT GAC GAC GAC GAT ATC TT-3' (SEQ ID NO:7) and its complementary sequence 5'-AGC TAA GAT ATC GTC GTC GTC AAC GAA-3' (SEQ ID NO:8), were cloned into pET31F1mhCA2 which had been digested with Hind III. This plasmid was designated pA1.

Plasmid pA1 was digested with the restriction endonucleases Ssp I and BspE I and the resulting ends were made blunt by treatment with T4 DNA polymerase. The DNA fragment from the pA1 digest containing the T7-hCAII-cassette was subcloned into the Sca I restriction site of pBR322 (New England Biolabs) thus conferring tetracycline resistance, but not ampicillin resistance. The resulting plasmid was designated pBN1.

Construction of pBN4

The pA1 plasmid was opened at the Hind III site and the EcoR V site and the synthetic oligonucleotide, 5'-A GCT GAA TTC AAC GTT CTC GAG GAT -3' (SEQ ID NO:9) and its complementary sequence 5'-ATC CTC GAG AAC GTT GAA TTC-3' (SEQ ID NO:10), were cloned into the vector. The insertion of these oligonucleotides provides a T7-hCAII-cassette containing unique EcoR I and Xho I restriction sites at the carboxyl terminal of hCAII. The resulting plasmid was designated pA3.

The pBN1 vector was digested with EcoR I and the single stranded overhangs were filled in with Polymerase I Large (Klenow) Fragment. The linear plasmid with newly formed blunt ends was religated, thus destroying the EcoR I site. The resulting plasmid was designated pBN3.

Plasmid pA3 was digested with the restriction endonucleases, Xba I and BspE I. The DNA fragment from the pA3 digest containing the T7-hCAII-cassette was subcloned into the pBN1 vector which had been digested with Xba I and BspE I. The resulting vector was designated plasmid pBN4.

Example 2. Inoculum Preparation for hCA-Fusion Constructs

L-broth was sterilized in the autoclave at 121°C for 20 minutes on the liquid cycle setting. The glucose and tetracycline stocks were filter sterilized by passage of the solution through a 0.22 μm filter. Two 250 ml shake flasks were each charged with the following solutions:

50 ml L-broth (1.0 % tryptone, 1.0% NaCl, 0.5% yeast extract)

1.0 ml glucose stock (50 mg/ml)

150 μl tetracycline stock (5.0 mg/ml)

100 μl thawed inoculum of *E. coli* cells transfected with a vector coding for the desired hCA-fusion construct

The shake flasks were placed in an incubator shaker at 37°C, 200-220 rpm for 10-14 hours. The optical density (O.D.) of the cells in the resulting solutions was then measured at 540 nm. A 1:25 dilution was usually necessary to obtain a proper reading. One of the two shake flasks was then chosen for inoculating the next set of shake flasks. Three 500 ml shake flasks were each charged with the following sterilized solutions:

200 ml L-broth (1.0 % tryptone, 1.0% NaCl, 0.5% yeast extract);

4.0 ml glucose stock (50 mg/ml);

600 μl tetracycline stock (5.0 mg/ml);

1.0 ml inoculum from one of the first two shake flasks.

The three shake flasks were placed in an incubator shaker under the conditions described above and the cells were allowed to grow for 8-10 hours. The optical density of the resulting solutions was then measured at 540 nm (typically at a 1:25 dilution). All three shake flasks were then used to inoculate the fermentor.

Example 3. Fermentation of hCA-Fusion Construct on 60 L Scale

Fermentation media was added to the fermentor and the volume was adjusted to 45.0 L with distilled H_2O. The media contained the following: 1200.0 g Case amino acids; 300.0 g Yeast extract; 30.0 g NaCl; and 0.10 ml Antifoam. The fermentor was sterilized at 121°C for 25 minutes. The fermentor was cooled to 37°C. Before inoculation, the following solutions were added to the fermentor:

glucose (480.0 g in 800.0 ml H_2O)

magnesium (120.0 g $MgSO_4 \cdot HO$ in 250.0 ml H_2O)

phosphates (120.0 g K_2HPO_4 & 465.0 g KH_2PO4 in 3.0 L H_2O)

tetracycline (0.90 g tetracycline·HCl in 30.0 ml 95% EtOH & 20.0 ml H_2O)

mineral mix (Dissolved in 490.0 ml H_2O & 10.0 ml concentrated HCl):

3.6 g $FeSO_4 \cdot 7H_2O$
3.6 g $CaCl_2 \cdot 2H_2O$
0.90 g $MnSO_4$
0.90 g $AlCl_3 \cdot 6H_2O$
0.09 g $CuCl_2 \cdot 2H_2O$
0.18 g Molybdic Acid
0.36 g $CoCl_2 \cdot 6H_2O$

All of the above solutions were sterilized for 20 minutes in the liquid cycle in an autoclave except for the tetracycline and mineral mix solutions. These were sterilized by passage through a .22 μm filter. At this point, the pH typically had dropped to approximately 6.5. If this had occurred, base (14.8 N Ammonium

hydroxide) was added to adjust the pH to 6.8. After the pH reached 6.8, 600.0 ml inoculum was added to the fermentor. The following parameters were monitored at time zero and throughout the fermentation: Glucose concentration (maintained at about 2-5 g/L); Optical Density; pH (6.8 is near optimal); Dissolved Oxygen (40% is near optimal); and Agitation. The temperature was maintained at 37°C throughout the fermentation. Air intake was 40 L/min at the beginning of the fermentation. The initial dissolved oxygen concentration was 90% but quickly dropped to 40%. It was maintained at this level via increased agitation and oxygen supplementation throughout the fermentation. When oxygen supplementation was started, the air inlet was reduced to 20 L/min. The initial glucose concentration was approximately 9 g/L but dropped to 5 g/L after about six hours. Once the glucose concentration dropped to this level, a glucose feed (70% w/v glucose) was used to maintain the glucose concentration at 5 g/L.

When the fermentation had proceeded to the point where an O.D. of 15-20 was measured, the media feed was started. The media feed consisted of 1200.0 g case amino acids and 300.0 g yeast extract dissolved in 5.0 L distilled H_2O and sterilized for 20 minutes on liquid cycle. The media feed was added to the fermentor over 1.0-1.5 hours. When fermentation had produced an O.D. of 30.0, the fermentation was induced by adding the following solutions to the fermentor: isopropylthiogalactoside (IPTG; 28.8 g in 200 ml distilled H_2O); $ZnCl_2$ (0.818g in 50 ml distilled H_2O with one drop of 6N HCl). The IPTG solution was filter sterilized through 0.22 μm filter. The $ZnCl_2$ solution was sterilized for 20 minutes using the liquid cycle in the autoclave. After the addition, the concentration of IPTG in the fermentor was 2.0 mM and the concentration of $ZnCl_2$ was 100 μM. A feed of a mixture of amino acids

was then started at this point. The amino acid feed consisted of 225.0 g L-serine; 75.0 g L-tyrosine; 74.0 g L-tryptophan; 75.0 g L-phenylalanine; 75.0 g L-proline; and 75.0 g L-histidine; and was dissolved in a mixture of 1.5 L H_2O and 500 ml concentrated HCl. The amino acid feed was sterile filtered through a 0.22 μm filter prior to addition to the fermentation. Induction was allowed to continue for 2.0 hours at which point the fermentation broth was transferred to a harvest tank and chilled to approximately 5-10°C. The fermentation typically yielded between 6.5 to 9.5 kg of wet cell paste (dry cell weight of about 1.0-1.5 kg).

Example 4. Harvest of 60 L Fermentation

The cell suspension from the fermentor as desribed above was concentrated over a tangential crossflow membrane to a volume of 10 L. The concentrated cell suspension was diafiltered and washed with 30 L of a cold wash buffer containing 50 mM Tris-SO4 pH 7.8, 1.0 mM EDTA, and 0.10 mM phenylmethylsulfonyl fluoride (PMSF). The cell suspension was then concentrated to 8 L. The concentration and washing of the cell suspension typically required 6-10 hours. At this point, the concentrated cell suspension (cell paste) may be bagged and frozen for later processing or transferred to homogenizer holding tanks for cell lysis.

Example 5. Cell Lysis

The cell paste obtained from the 60 L fermentor was diluted to 32 L in cold wash buffer (see above) and the resulting cell suspension was chilled to 5-10°C. The chilled cell suspension was homogenized at 12,000 psi with a Galin high pressure homogenizer. The homogenized cell paste was passed through a heat exchanger to chill the lysate to 10°C and passed through the homogenizer a second time.

Example 6. Preparation of Thrombin Stock Solution

The thrombin used in fusion protein cleavage reactions was obtained from Calbiochem in the form of a lyophilized powder. This powder was solubilized prior to use by dissolution in MilliQ water to a concentration of 1 mg/ml as determined by specific activity.

Example 7. Production of PTH(1-34) (SEQ ID NO:22)

The preparation of the DNA segment coding for a single copy PTH(1-34) (SEQ ID NO:22) fusion protein was carried out by preparing an expression vector coding for a PTH-construct which included a DNA segment coding for the hCAII, an interlinking peptide, and PTH(1-34) (SEQ ID NO:22). The following oligos were obtained from Operon Technologies Inc, 1000 Atlantic Ave. Alameda CA 94501.

Oligo 1: 5' CCC AAG CTT CTG TTC GTG GTC CGC GTT CTG TTT CTG AAA (SEQ ID NO:11)

Oligo 2: 5' GAA ACA GAA CGC GGA CCA CGA ACA GAA GCT TGG G (SEQ ID NO:12)

Oligo 3: 5' TCC AGC TGA TGC ACA ACC TGG GTA AAC ACC TGA ACT (SEQ ID NO:13)

Oligo 4: 5' AGG TGT TTA CCC AGG TTG TGC ATC AGC TGG ATT TCA (SEQ ID NO:14)

Oligo 5: 5' CTA TGG AAC GTG TTG AAT GGC TGC GTA AAA AAC TGC A (SEQ ID NO:15)

Oligo 6: 5' TTT TTT ACG CAG CCA TTC AAC ACG TTC CAT AGA GTT C (SEQ ID NO:16)

Oligo 7: 5' GGA CGT TCA CAA CTT CTA AGA TAT CCG G (SEQ ID NO:17)

Oligo 8: 5' CCG GAT ATC TTA GAA GTT GTG AAC GTC CTG CAG (SEQ ID NO:18)

The eight synthetic DNA oligos were obtained and the complementary strands phosphorylated and annealed. The four double stranded fragments were used to prepare a DNA fragment coding for the interpeptide linker followed by PTH(1-34) (SEQ ID NO:22). This DNA fragment

was digested and inserted into pA1 using the restriction sites Hind III and EcoR V creating the vector pA1:PTH(1-34).

The expression cassette was removed from pA1:PTH(1-34) by digestion with the restriction endonucleases Xba I and BspE I and inserted into the vector pBN1 which had been digested with the same restriction endonucleces. This final vector was desiganted pBN1:PTH(1-34).

E. coli cells transformed with the vector pBN1:PTH(1-34) containing DNA coding for an hCA-PTH fusion protein were prepared, cultured and lysed according to the procedures described above. The hCA-PTH fusion protein included hCAII linked to an amino acid sequence corresponding to PTH(1-34) (SEQ ID NO:22) through a thrombin cleavage site (Gly-Pro-Arg) immediately adjacent the N-terminus of the PTH(1-34) (SEQ ID NO:22) sequence. The hCA-PTH fusion protein was expressed as inclusion bodies in *E. coli* BL21 $F^- ompT\ r_B^- m_B^-$ (DE3) cells.

Inclusion Body Purification

The cell lysate from the 60 L fermention run was centrifuged in a continuous flow rotor at 20,000 g at 400 ml/min. The pellet contained 300-500 g of crude inclusion bodies. The crude hCA-PTH inclusion bodies was suspended in 2M citic acid at a concentration of 60 mg solids/ml (4-8 liters) and sonicated at 70% power with a 50% pulse rate. The solubilized inclusion bodies were clarified by centrifugation at 20,000 g.

The hCA-PTH fusion protein in the supernatant was precipitated by the addition of 10 N NaOH until the pH reaches 4-5. The precipitated hCA-PTH fusion protein was collected by centrifugation at 20,000 g. The precipitated hCA-PTH fusion protein was then washed with 4 to 8 liters of 5% acetic acid/45% EtOH in Milli-Q water and the precipitated hCA-PTH fusion protein was again collected by centrifugation. The hCA-PTH fusion

protein was then washed twice with 4 to 8 liters of 100 mM EDTA and once with 4-8 liters of Milli Q water and the protein collected by centrifugation after each wash.

Thrombin Cleavage of PTH (1-34)

A solution containing 200 ml of 50mM NaOH and 0.25% N-lauroyl sarcosine is added to a bottle containing a pellet of the hCA-PTH fusion protein. The bottles are placed into 37°C water bath to warm pellets (typically for 2-5 minutes) to aid in the solubilization of the pellet. The material is homogenized until all large pieces are disaggregated. The pH is readjusted 11.6 to 11.9 with a solution containing 50mM NaOH and 0.25% N-lauroyl sarcosine. The resulting solution is sonicated until all of the hCA-PTH fusion protein pellet has dissolved. Typically sonication is for 2 minutes at power 10, pulser on, 70% duty cycle (smaller batches require less time). The protein concentration of the reaction solution is determined by absorbance at 280 nm. Ideally the concentration should be 6-7 mg/ml (8-10 liters for a 60 l fermentation). If the concentration is greater than 9 mg/ml, the solution is diluted to 6-7 mg/ml with the 50mM NaOH/0.25% N-lauroyl sarcosine solution.

The protein solution is stirred vigorously and the pH is adjust to 8-8.2 with 1M Tris-HCl. If solution is hazy, it is filtered through glass fiber filters. The solution is then filtered through 0.45 μm cellulose acetate membrane and the protein concentration is determined by measuring the absorbance at 280 nm. Solid NaCl is added to a final concentration of 250mM. A thrombin stock solution 1 mg/ml is added to produce a protein to enzyme ratio of 5000:1 w/w. The solution is then stirred in a water bath plate at 37°C. The reaction is monitored by HPLC on a C8 column eluted with a gradient from Buffer A (95% water, 5% acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA)) to Buffer B (95%

ACN, 5% water, .1% TFA). The reaction is stopped by the addition of PMSF. Typically the reaction is complete as determined by the loss of the peak corresponding to the starting material in 46-48 hours. The PMSF is dissolved into 95% EtOH, and added directly to the reaction mixture with vigorous stirring. The thrombin cleavage procedure yields 5.1 g of PTH(1-34) (SEQ ID NO:22) from 60 g of hCA-PTH fusion protein.

Citric Acid Precipitation of Thrombin Cut PTH Material

Both the resulting human carbonic anhydrase (from the cleavage of the hCA-PTH fusion protein) and remaining unhydrolized fusion protein are precipitated from the solution by adding sufficient citric acid to achieve 150mM, leaving the desired peptide (PTH 1-34) (SEQ ID NO:22) in solution. The precipitated material is removed by centrifugation. The supernatant is filtered through 0.45 μm filter and frozen at -80°C or desalted directly on a C8 column. The yield from the peptide from the step is 85%.

If desired the PTH(1-34) (SEQ ID NO:22) may be further purified by chromatography on preparative C8 column using a gradient from 10-70% Buffer B of aqueous acetonitrile/ trifluoroacetic acid (ACN/TFA) buffers, where Buffer A:0.1% TFA, 95% water, 5% acetonitrile, Buffer B:0.1% TFA, 5% water, 95% acetonitrile. Alternatively, the PTH(1-34) (SEQ ID NO:22) sample may be purified on the preparative C8 column eluting first with a gradient from 10-70% Buffer B; where Buffer A is 5mM HCl, 95% water, 5% acetonitrile and Buffer B. The PTH(1-34) (SEQ ID NO:22) is desalted with a gradient where Buffer A 10mM acetic acid, and Buffer B is 10 mM acetic acid in 50% aqueous ethanol.

Example 8. Production of GLP1(7-36)-NH$_2$ (SEQ ID NO:27)

Six synthetic DNA oligos (1-6) were phosphorylated and the complementary strands annealed. Oligopair 1&2

were inserted into pUC19 (commercially available from New England Biolabs. MA) between Hind III and Kpn I. Oligopair 3&4 was then inserted into this vector between Kpn I and EcoR I yielding a vector named pUC19:GLP(7-22). Oligopair 5&6 was ligated into pBluescript® IISK+ (commercially available from Stratagene) between Hind III and EcoR I yielding the vector PB5:GLP(21-34)AFA. The fragment containing oligopair 5&6 was inserted into pUC19:GLP(7-22) by digesting pUC19:GLP(7-22) with Nar 1 and EcoR I and digesting PB5:GLP(21-34)AFA with Cla I and EcoR I and ligating the fragment from PB5:GLP(21-34)AFA containing oligos 5&6 into the truncated pUC19:GLP(7-22) to yield pUC19:GLP(7-34)AFA. pUC19:GLP(7-34)AFA was digested with Hind III and the fragment containing the sequence coding for the peptide interlinker followed by GLP1(7-34)-Ala-Phe-Ala (SEQ ID NO:30) ("GLP1(7-34)AFA") was transferred to pA4 to yield vector pA4:GLP(7-34)AFA. The hCAII gene in pA4 had been mutated to change the methionine-240 to a cysteine. The expression cassette was then digested from pA4:GLP(7-34)AFA with the restriction endonucleases Xba I and BspE I and inserted into the vector pBN1 digested with the same restriction endonucleases. The resulting vector was designated pBN6:GLP(7-34)AFA.

E. coli cells transformed with the vector pBN6:GLP1(7-34)AFA containing DNA coding for an hCA-GLP1(7-34)-Ala-Phe-Ala ("hCA-GLP1-AFA") fusion protein were prepared, cultured and lysed according to the procedures described above. The hCA-GLP1-AFA fusion protein included hCAII linked to an amino acid sequence corresponding to GLP1(7-34)-Ala-Phe-Ala (SEQ ID NO:30) through a methionine residue located immediately adjacent the N-terminus of the GLP1(7-34)-Ala-Phe-Ala (SEQ ID NO:30) sequence. The hCA-GLP1-AFA fusion protein was expressed as inclusion bodies in *E. coli* BL21 F^- ompT r_B^- m_B^- (DE3) cells.

Inclusion Body Purification

The cell lysate was centrifuged in a continuous flow rotor at 20,000 g at 400 ml/min to produce a pellet containing 300-500 g of crude inclusion bodies. The crude hCA-GLP1-AFA inclusion bodies were suspended in lysis buffer at a concentration of 25% solids (1.2-2.0 L) and homogenized in the Galin homogenizer at 13,000 psi. The homogenized inclusion bodies were collected by centrifugation at 20,000 g, resuspended in a solution of 100 mM Na Acetate pH 4.5 and homogenized a second time. The homogenized inclusion bodies were then suspended in distilled water. The resulting suspension is homogenized and the inclusion bodies collected by centrifugation. The purified inclusion bodies are dissolved in 1.5 M citric acid and clarified by centrifugation. The hCA-GLP1-AFA fusion protein was precipitated by the addition of NaOH until the pH is 4-5 and isolated by centrifugation.

Cyanogen Bromide Cleavage of hCA-GLP1-AFA

The hCA-GLP1-AFA inclusion bodies were dissolved in a solution of 2.0 M Citric Acid at a concentration of about 35-45 mg/ml. The solution was adjusted pH to 1.0 with conc. HCl and 0.2 g cyanogen bromide (CNBr) per g of total protein was added. The protein concentration was determined by HPLC. The protein solution was sparged with pure Argon before addition of the CNBr. The cleavage reaction was allowed to proceed under an Argon atmosphere for 4-5 hours at which time residual CNBr was neutralized by adding 2.2g D,L-methionine per g of CNBr added followed by stirring for 30 minutes.

Purification of GLP1(7-34)-AFA (SEQ ID NO:30)

The solution from the CNBr cleavage reaction was diluted 1:1 with a solution of 10% sodium sulfate (w/v) in water to precipitate all protein and peptide from the solution. The sample was centrifuged at 20,000 x g for

15 minutes and the supernatant decanted. Distilled water (200 ml) was added to each centrifuge vial. The resulting mixture was homogenized and centrifuged for 15 min. at 20,000 x g. The supernatant was decanted and a volume of 50% acetonitrile (ACN) was added to the pellet material. The sample was homogenized, stirred for 10-15 minutes, and finally centrifuged at 20,000 x rpm for 15 minutes. The extraction with the 50% acetonitrile solution was repeated two to three times to achieve a 90-95% recovery of GLP1(7-34)AFA (SEQ ID NO:30) in the acetontrile solution. The extraction may also be carried out with 50% n-propanol.

Batch loading of 50% ACN extraction onto DOWEX-1

DOWEX-1 (25 g for each 250 mg of GLP1(7-34)-AFA (SEQ ID NO:30) is prewashed with 10% ACN solution prior to being added to the 50% ACN extraction solution. The DOWEX-1 is loaded in a batch fashion and the suspension stirred for 20 minutes. The DOWEX-1 is then filtered away from the GLP1(7-34)-AFA (SEQ ID NO:30) solution.

Low Pressure C8 chromatography

The 50% ACN solution is diluted to 12.5% ACN with distilled H_2O to facilitate the binding of GLP1(7-34)-AFA (SEQ ID NO:30) to the C8 resin. The peptide solution is loaded onto the C8 resin in an low pressure column and the column is washed with 5 column volumes of 30% (v/v) ACN solution. The GLP1(7-34)-AFA (SEQ ID NO:30) is then eluted with a solution of 50% ACN in distilled H_2O. The 50% ACN solution is lyophilized to obtain the GLP1(7-34)-AFA (SEQ ID NO:30).

Transpeptidation of GLP1(7-34)AFA (SEQ ID NO:30) into GLP1(7-36)NH_2 (SEQ ID NO:27)

GLP1(7-34)AFA (SEQ ID NO:30) (1 mmole) is dissolved in a 50/50 (v/v) solution of DMF/H_2O which includes 5 mM $CaCl_2$ and 400 mM Gly-Arg-NH_2 and the solution pH is adjusted to 8.0 with 5 M NaOH. Two mg of TPCK-treated

trypsin is then added. The transpeptidation reaction which produce GLP1(7-36)NH_2 (SEQ ID NO:27) is complete in 2-3 hours at room temperature.

If desired the GLP1(7-36)-NH_2 (SEQ ID NO:27) may be further purified by chromatography on preparative C8 column using a gradient of aqueous acetonitrile/trifluoroacetic acid (ACN/TFA) buffers.

Example 9. Production of GRF(1-44)NH_2 (SEQ ID NO:20)

Eight oligonucleotides containing segments of the linker and the peptide were phosphorylated and complementary oligonucleotide pairs 1&2, 3&4, 5&6, and 7&8 were annealed. Oligonucleotide pairs 1&2 and 3&4 were simultaneously ligated into the pTZ19R vector (commercially available from Pharmacia Biotech Inc., NJ) between the Hind III and Sal I sites to yield pTZ:GRF(1-29). Oligonucleotide pairs 5&6 and 7&8 were simultaneously ligated into a separate pTZ19R vector between the Sal I and EcoR I sites to yield pTZ:GRF(29-44)A. The gene fragments were cloned adjacent to each other in a single vector by digesting pTZ:GRF(1-29) and pTZ:GRF(29-44)A with the restriction endonucleases Xmn I and Sal I, isolating the 1.9 kb band from the pTZ:GRF(1-29) vector and the 0.9 kb band of the pTZ:GRF(29-44)A vector and ligating them together to yield the vector pTZ:GRF(1-44)A.

The three Asn-Gly sites in hCAII (located at positions 10-11, 61-62 and 230-231) were changed to Gln10-Gly11, Gln61-Gly62 and Asn230-Ala231 by site mutagenesis of specific codons in plasmid pA1. All three mutations were combined to create plasmid pA2.

The pA2 vector was digested with EcoR V. pTZ:GRF(1-44)A was digested with Dra I and EcoR V. The fragment containing the GRF gene and the linearized pA2 plasmid were ligated together to yield pBN2:GRF(1-44)A.

E. coli cells transformed with vector pBN2:GRF(1-44)A containing DNA coding for an hCA-GRF fusion protein

were prepared, cultured and lysed according to the procedures described above. The hCA-GRF fusion protein included hCAII linked to an amino acid sequence corresponding to GRF(1-44)-Ala (SEQ ID NO:31) through an interconnecting peptide. The interconnecting peptide included a thrombin cleavage site and an enterokinase cleavage site. The enterokinase cleavage site (Asp-Asp-Asp-Asp-Lys) (SEQ ID NO:2) was positioned immediately adjacent the N-terminus of the GRF(1-44)-Ala (SEQ ID NO:31) sequence. The thrombin cleavage site (Gly-Pro-Arg) was located so that after treatment with thrombin, the hCA-GRF fusion protein produced a peptide fragment having an 8 amino acid sequence (Ala-Met-Val-Asp-Asp-Asp-Asp-Lys) (SEQ ID NO:19) connected to the N-terminus of the GRF(1-44)-Ala (SEQ ID NO:31) sequence. The hCA-GRF fusion protein was expressed as inclusion bodies in *E. coli* BL21 F^- ompT $r_B^- m_B^-$ (DE3) cells.

Inclusion Body Purification

The cell lysate was centrifuged in a continuous flow rotor at 20,000 g at 400 ml/min to produce a pellet containing 300-500 g of crude inclusion bodies. The crude hCA-GRF-Ala inclusion bodies were suspended in 2M citic acid at a concentration of 60 mg solids/ml (5-9 L) and sonicated at 70% power with a 50% pulse rate. The solubilized inclusion bodies were clarified by centrifugation at 20,000 g. The hCA-GRF-Ala fusion protein in the supernatant was precipitated by the addition of 10 N NaOH until the pH reaches 4-5. The precipitated hCA-GRF-Ala fusion protein was then collected by centrifugation at 20,000 g. After washing with 5% acetic acid/ 45% ethanol in Milli-Q H_2O, the precipitated hCA-GRF-Ala fusion protein was again collected by centrifugation. The precipitated hCA-GRF-Ala fusion protein was then resuspended twice in 100 mM EDTA and the fusion protein was collected by centrifugation. The washed, precipitated hCA-GRF-Ala

fusion protein was suspended in distilled H_2O. The suspension was homogenized and the hCA-GRF-Ala fusion protein was collected by centrifugation.

Thrombin Cleavage of the hCA-GRF Fusion Protein

A 200 ml portion of a solution containing 50mM NaOH and 0.25% N-lauroyl sarcosine was added to a bottle containing a pellet of the hCA-GRF-Ala fusion protein. The bottle was placed into 37°C water bath to warm the pellet. The NaOH solution was swirled around in the bottle to remove all solid material from the bottle sides. The slurry was then poured into a suitable size beaker. The material was homogenized until all large pieces are disaggregated and the pH was monitored. The pH was readjusted to between 11.6 and 11.9 with a solution of 50mM NaOH and 0.25% N-lauroyl sarcosine. The solution was sonicated until all of the inclusion body pellet has dissolved and solution is clear. Sonication times varied according to size of batch. Typically anything over 1 liter was sonicated for 2 minutes, power 10, pulser on, 70% duty cycle (smaller batches required less time).

The protein concentration of the reaction solution was measured. If concentration was greater than 9 mg/ml, the solution was diluted with the 50mM NaOH/0.25% N-lauroyl sarcosine solution to a protein concentration of 6-7 mg/ml. The protein solution was stirred vigorously, and the pH adjusted to 8-8.2 with 1M Tris-HCl. If the solution was slightly hazy at this point, the solution was clarified by filtering through glass fiber filters. The clarified solution was sterile filtered through a 0.45 μm cellulose acetate membrane. The protein concentration was determined by measuring the absorbance at 280 nm. Sufficient NaCl to produce a 250mM NaCl concentration was added with vigorous stirring.

Sufficient thrombin stock solution (1 mg/ml) to

produce a protein to enzyme ratio of 5000:1 was added and the resulting mixture stirred in a water bath on top of stir plate at 37°C. The reaction was monitored by HPLC on a C8 column eluted with a gradient of H_2O/ACN/TFA buffers. The reaction was stopped by the addition of PMSF to a final concentration of 0.1mM when the fusion construct peak was essentially gone (usually 46-48 hours). The PMSF was dissolved into 95% EtOH and added directly to the reaction mixture with vigorous stirring.

Citric Acid Precipitation of the hCA-Fragment

The human carbonic anhydrase fragment from the cleavage of the fusion protein and any residual uncut fusion protein were removed from solution by precipitation with 90mM citric acid, leaving the intermediate peptide, Ldr-GRF(1-44)-Ala, in solution.

The volume of the thrombin cut solution was measured in a graduated cylinder. A stock solution of 1M citric acid was added to final concentration of 90mM citrate. The addition was made slowly while with vigorously stirring the solution. The aggregated material was centrifuged and the supernatant was filtered through a 0.45 μm filter. The filtered supernatant was either frozen at -80°C or run directly onto a C8 column for desalting.

The pellets from the precipitation were washed with a solution of 90 mM citrate pH 3.0-3.1. Between 150-200 ml of citrate solution was added to each pellet of 1-2 grams of hCA. The pellets were homogenized and the supernatant from the wash was filtered through a 0.45 μm filter and saved. The washed pellets were discarded. The supernatant was either frozen at -80°C or desalted directly on a C8 column.

C8 Desalting of Ldr-GRF(1-44)-Ala

The 90 mM citric acid solution of Ldr-GRF(1-44)-Ala was loaded directly onto a preparative C8 column. The

column was eluted with a gradient of aqueous ethanol/acetic acid buffers. The fractions containing the Ldr-GRF(1-44)-Ala were collected (typically 90-95% pure) and the solution concentrated by evaporation to 21-25 mg/ml. The resulting solution was stored at -80°C prior to transpeptidation.

Transpeptidation of Ldr-GRF(1-44)-Ala

The concentration Ldr-GRF(1-44)-Ala peptide was determined by C8 HPLC using 1 mM GRF standard. The peptide solution was diluted to a concentration of 3 mM Ldr-GRF(1-44)-Ala (18.0 mg/ml) with water. EDTA and sodium phosphate were added (to 5 mM; 1.9 mg/ml and 25 mM; 3.6 mg/ml respectively) and the pH was adjusted to 6.0 with 1 M NaOH. 2-Nitrophenylglycinamide [ONPGA] (250 mM; 89.0 m g/ml) was added and the pH was again adjusted to 6.0 with 1 M NaOH. Carboxypeptidase-Y ("CPD-Y"; 2 µl/ml) was the added and stirring was in the dark at 35-40 °C. The extent of reaction to produce Ldr-GRF-ONPGA was monitored by HPLC. After about 1-2 hours the reaction was stopped by the addition of acetonitrile to a final concentration of 15% v/v.

The sample from the CPD-Y transpeptidation was loaded onto a C8 HPLC column and rinsed with a 20% ethanol, 10 mM sodium phosphate, pH 6.8 buffer to remove unreacted ONPGA. The column was then eluted with a 50% EtOH, 10 mM Sodium Phosphate, pH 6.8 buffer and the Ldr-GRF-ONPGA peak was collected.

Photolysis of Ldr-GRF-ONPGA

The concentration of the Ldr-GRF-ONPGA peptide was determined by C8 HPLC using 1 mM GRF standard and the concentration was diluted to 1 mM Ldr-GRF-ONPGA (6.0 mg/ml) with 50% ethanol. Sodium bisulfite (to 5 mM) and sodium benzoate (to 50 mM) were added and the solution is adjusted to pH 9.5 with 1 M NaOH. After the solution has been purged with Argon for 10-15 minutes, the

mixture was irradiated (wavelength >305 nm; 200-210 W medium pressure mercury lamp) while being maintained in a 20-25 °C circulating water bath. During the reaction a constant flow of Argon was maintained and the reaction was constantly stirred. The reaction was monitored by HPLC. After about 1-2 hours reaction was stopped by lowering the pH to 5.5 with acetic acid.

Enterokinase Cleavage Protocol For Ldr-GRF-NH2

The solution of Ldr-GRF-NH2 from the photolysis reaction was diluted 1:5 with water (1.0 mg/ml peptide). Triton X-100 was added to a final concentration of 0.1%. Succinic acid and calcium chloride were added to produce concentrations of 50 mM (5.9 mg/ml) and 2 mM (0.3 mg/ml) respectively and the solution pH was adjusted to 5.5 with 10 M NaOH. After the solution was filtered through a 0.45 μm membrane, 5.0 mg/ml Dowex 1 resin was added. The Dowex 1 resin, an anion exchange resin, was added to the reaction to bind the peptide containing the Asp-Asp-Asp-Asp (SEQ ID NO:32) sequence. This was found to increase both the rate of the reaction and the overall yield, as well as reduce the concentration of enterokinase required while improving substrate specificity. A 1:3000 ratio of enterokinase enzyme (1 ul per 10 mg peptide) was added and the reaction was maintained in a 35-40 °C water bath with constant stirring. After 20-24 hours, the cleavage reaction which converts the Ldr-GRF-NH2 into GRF(1-44)-NH2 (SEQ ID NO:20) reached 70-80% completion. The reaction mixture was filtered to remove the Dowex 1 and the reaction was stopped by the addition of acetonitrile to a final concentration of 15%. The sample was stored at -80°C until ready for purification. If desired, purification of the GRF(1-44)-NH2 (SEQ ID NO:20) product may be carried out by preparative HPLC using a C8 column.

Example 10. Production of Peptide via Expression of a Soluble hCA Fusion Protein Construct

E.coli cells transformed with a vector containing DNA coding for a hCA-9AA fusion protein may be prepared and cultured according to the procedures described above. The hCA-9AA fusion protein includes hCAII linked to the amino acid sequence Thr-Asn-Thr-Gly-Ser-Gly-Thr-Pro-Ala ("9AA") (SEQ ID NO:33) through an interconnecting peptide. The interconnecting peptide includes an enterokinase cleavage site positioned immediate adjacent the N-terminus of the 9AA (SEQ ID NO:33) sequence. The hCA-9AA fusion protein may be expressed as a soluble protein in *E. coli* BL21 F' ompT r_B m $_B$ (DE3) cells.

The cell paste from the fermentor is diluted with cold wash buffer and chilled to 5-10°C. The chilled cell suspension is homogenized at 12,000 psi with Galin high pressure homogenizer. The cell paste is passed through a heat exchanger and chilled to 10°C prior to a second pass through the homogenizer. A PMSF stock solution in 95% ethanol is added to a final concentration of 0.05 mM PMSF. The lysate is cooled to about 8°C, passed through the homogenizer at 12,000 psi and drained into a holding tank. After passing the lysate through the homogenizer at 12,000 psi yet another time, a 5% v/v pH 7.5 solution of polyethyleneimine (PEI) may be added to a total concentration of 0.35% PEI. The resulting mixture is stirred for 20 minutes and clarified by Westfalia at 20,000 x g. The clarified solution is filtered through a 0.22 micro μm Pall filter system.

A affinity column including p-aminomethylbenzene-sulfonamide-agarose resin is equilibrated with a 0.1M Tris-SO_4 pH 8.7 solution. The clarified supernatant from the PEI precipitation is loaded onto the column and the column is eluted with a series of Tris-SO_4 buffer solutions. The column fractions containing the fusion

protein construct are pooled. The protein construct is precipitated from the pooled fractions by lowering the pH of the solution to 4.0 with glacial acetic acid. The precipitated fusion protein construct may be isolated by centrifugation and lyophilized.

The precipitated fusion protein construct is dissolved in a ph 8.0 buffer solution containing 50mM Tris and 1mM $CaCl_2$. The mixture is sonicated until all the fusion protein construct has been dissolved. After the solution was filtered through a 0.45 um membrane, 5.0 mg/ml Dowex 1 resin was added. A 1:3000 ratio of enterokinase enzyme (1 ul per 10 mg peptide) was added and the reaction was maintained in a 35-40 °C water bath with constant stirring for 20-24 hours. The reaction mixture was filtered to remove the Dowex 1 and the reaction was stopped by the addition of acetonitrile to a final concentration of 15%. A 3:1 volume of 15% (v/v) acetic acid is added to the solution and the resulting mixture is stirred for 30 minutes, thereby precipitating the peptide fragment which includes hCA. The hCA fragment is removed by centrifugation yielding a supernatant which contains the 9AA (SEQ ID NO:33) peptide. The supernatant may be desalted on a C-8 column and stored at -80°C.

The invention has been described with reference to various specific and preferred embodiments and techniques. However, It will be apparent to one of ordinary skill in the art that many variations and modifications may be made while remaining within the spirit and scope of the invention.

All publications and patent applications in this specification are indicative of the level of ordinary skill in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated by reference.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: BioNebraska, Inc.

(ii) TITLE OF THE INVENTION: PRODUCTION OF PEPTIDES USING RECOMBINANT FUSION PROTEIN CONSTRICTS

(iii) NUMBER OF SEQUENCES: 33

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Merchant & Gould
(B) STREET: 3100 Norwest Center, 90 S. 7th Street
(C) CITY: Minneapolis
(D) STATE: MN
(E) COUNTRY: U.S.A.
(F) ZIP: 55402

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 07-DEC-1995
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/350,530
(B) FILING DATE: 07-DEC-1994

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Carter, Charles G
(B) REGISTRATION NUMBER: 35,093
(C) REFERENCE/DOCKET NUMBER: 8648.45USWO

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 612/332-5300
(B) TELEFAX: 612/332-9081
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...159
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GAC GAC GAC GAT AAA                                              15
Asp Asp Asp Asp Lys
 1               5
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Asp Asp Asp Asp Lys
 1               5
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...12
(D) OTHER INFORMATION:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 1...0
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ATT GAA GGA AGA                                                  12
Ile Glu Gly Arg
 1
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ile Glu Gly Arg
1

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...24
(D) OTHER INFORMATION:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 1...0
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CAT CCT TTT CAT CTG CTG GTT TAT 24
His Pro Phe His Leu Leu Val Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

His Pro Phe His Leu Leu Val Tyr
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGCTTTCGTT GACGACGACG ATATCTT 27

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGCTAAGATA TCGTCGTCGT CAACGAA 27

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGCTGAATTC AACGTTCTCG AGGAT 25

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATCCTCGAGA ACGTTGAATT C 21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCCAAGCTTC TGTTCGTGGT CCGCGTTCTG TTTCTGAAA 39

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GAAACAGAAC GCGGACCACG AACAGAAGCT TGGG 34

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TCCAGCTGAT GCACAACCTG GGTAAACACC TGAACT 36

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGGTGTTTAC CCAGGTTGTG CATCAGCTGG ATTTCA 36

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTATGGAACG TGTTGAATGG CTGCGTAAAA AACTGCA 37

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TTTTTTACGC AGCCATTCAA CACGTTCCAT AGAGTTC 37

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGACGTTCAC AACTTCTAAG ATATCCGG 28

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE:
(vi) ORIGINAL SOURCE:
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCGGATATCT TAGAAGTTGT GAACGTCCTG CAG 33

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: N-terminal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ala Met Val Asp Asp Asp Asp Lys
1 5

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Tyr Ala Asp Ala Ile Phe Thr Asn Ser Tyr Arg Lys Val Leu Gly Gln
1 5 10 15
Leu Ser Ala Arg Lys Leu Leu Gln Asp Ile Met Ser Arg Gln Gln Gly
20 25 30
Glu Ser Asn Gln Glu Arg Gly Ala Arg Ala Arg Leu
35 40

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys
20 25

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
Ser Val Ser Glu Ile Gln Leu Met His Asn Leu Gly Lys His Leu Asn
 1               5                  10                  15
Ser Met Glu Arg Val Glu Trp Leu Arg Lys Lys Leu Gln Asp Val His
            20                  25                  30
Asn Phe
```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Tyr Ala Asp Ala Ile Phe Thr Asn Ser Tyr Arg Lys Val Leu Gly Gln
 1               5                  10                  15
Leu Ser Ala Arg Lys Leu Leu Gln Asp Ile Met Ser Arg Gln Gln Gly
            20                  25                  30
Glu Ser Asn Gln Glu Arg Gly Ala Arg
        35                  40
```

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
His Asp Glu Phe Glu Arg His Ala Glu Gly Thr Phe Thr Ser Asp Val
 1               5                  10                  15
Ser Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu
            20                  25                  30
Val Lys Gly Arg
        35
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 84 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
Ser Val Ser Glu Ile Gln Leu Met His Asn Leu Gly Lys His Leu Asn
 1               5                  10                  15
Ser Met Glu Arg Val Glu Trp Leu Arg Lys Lys Leu Gln Asp Val His
            20                  25                  30
Asn Phe Val Ala Leu Gly Ala Pro Leu Ala Pro Arg Asp Ala Gly Ser
        35                  40                  45
Gln Arg Pro Arg Lys Lys Glu Asp Asn Val Leu Val Glu Ser His Glu
    50                  55                  60
Lys Ser Leu Gly Glu Ala Asp Lys Ala Asp Val Asn Val Leu Thr Lys
65                  70                  75                  80
Ala Lys Ser Gln
```

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```
His Asp Glu Phe Glu Arg His Ala Glu Gly Thr Phe Thr Ser Asp Val
 1               5                  10                  15
Ser Ser Tyr Leu Glu Gly Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu
            20                  25                  30
Val Lys
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

```
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Cys
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```
Ser Val Ser Glu Ile Gly Leu Met His Asn Leu Gly Lys His Leu Asn
 1               5                  10                  15
Ser Met Glu Arg Val Glu Trp Leu Arg Lys Lys Leu Gln Asp Val His
            20                  25                  30
Asn Phe Val Ala Leu Gly
        35
```

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

```
His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
 1               5                  10                  15
```

```
Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Ala Phe Ala
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Tyr Ala Asp Ala Ile Phe Thr Asn Ser Tyr Arg Lys Val Leu Gly Gln
 1               5                  10                  15
Leu Ser Ala Arg Lys Leu Leu Gln Asp Ile Met Ser Arg Gln Gln Gly
            20                  25                  30
Glu Ser Asn Gln Glu Arg Gly Ala Arg Ala Arg Leu Ala
        35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Asp Asp Asp Asp
 1
```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Thr Asn Thr Gly Ser Gly Thr Pro Ala
 1               5
```

WHAT IS CLAIMED IS:

1. A method of producing a peptide comprising:
precipitating a fusion protein construct, which includes a carbonic anhydrase and a variable fused polypeptide.

2. The method of claim 1 comprising precipitating the fusion protein construct from a solution at a pH of about 3.2 to about 6.0.

3. The method of claim 2 wherein precipitating the fusion protein construct further comprises adding a salt to the solution.

4. The method of claim 3 wherein the salt includes a divalent metal cation.

5. The method of claim 2 wherein precipitating the fusion protein construct comprises adding an acid to the solution.

6. The method of claim 2 wherein precipitating the fusion protein construct comprises adding an base to the solution.

7. The method of claim 1 comprising precipitating the fusion protein construct from a solution which includes ammonium sulfate or sodium sulfate.

8. The method of claim 1 comprising precipitating the fusion protein construct from a solution which includes a chaotropic agent.

9. The method of claim 8 wherein the chaotropic agent comprises guanidine hydrochloride.

10. The method of claim 1 comprising precipitating the fusion protein construct from a solution which includes urea.

11. The method of claim 1 wherein isolating the fusion protein construct includes resolubilizing the precipitated fusion protein construct.

12. The method of claim 11 comprising resolubilizing the precipitated fusion protein construct in a solution having a pH of at least about 10.

13. The method of claim 11 comprising resolubilizing the precipitated fusion protein construct in a solution having a pH of no more than about 3.2.

14. The method of claim 11 comprising resolubilizing the precipitated fusion protein construct in a solution which includes a chaotropic agent or a detergent.

15. The method of claim 1 further comprising cleaving the fusion protein construct to produce a soluble carbonic anhydrase fragment and a soluble variable fused polypeptide fragment.

16. The method of claim 15 further comprising precipitating the carbonic anhydrase fragment.

17. The method of claim 15 further comprising contacting a solution of the carbonic anhydrase fragment with a support which includes a benzenesulfonamide compound or a acetazolamide compound.

18. The method of claim 1 wherein the variable fused polypeptide comprises an amino acid sequence

corresponding to a peptide selected from the group consisting of GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21), and PTH(1-34) (SEQ ID NO:22).

19. The method of claim 1 further comprising expressing the fusion protein construct as part of an inclusion body in a host cell.

20. The method of claim 19 further comprising dissolving the inclusion body to produce a solubilized fusion protein construct.

21. The method of claim 20 comprising dissolving the inclusion body in a solution which includes citric acid.

22. The method of claim 20 comprising dissolving the inclusion body in a solution which includes a chaotropic agent or a detergent.

23. The method of claim 20 comprising dissolving the inclusion body in a solution having a pH of at least about 10.

24. The method of claim 1 wherein the variable fused polypeptide includes at least 2 copies of a target peptide.

25. The method of claim 24 wherein the variable fused polypeptide includes an amino acid sequence having the formula:

$$-TargP-(CS2)-[\ -(Ln1)_n-(CS1)_m-TargP-(CS2)-\ ]_r$$

wherein the -CS1- and -CS2- are cleavage sites, the -(Ln1)- is a linking sequence, the -TargP- is the target peptide, n and m are 0 or 1, and r is an integer from 1 to about 150.

26. The method of claim 25 wherein the target peptide includes an amino acid sequence corresponding to a peptide selected from the group consisting of GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22).

27. The method of claim 1 wherein the carbonic anhydrase is linked to the variable fused polypeptide by a cleavage site.

28. The method of claim 27 wherein the cleavage site is a chemical cleavage site.

29. The method of claim 27 wherein the cleavage site is an enzymatic cleavage site recognized by an endopeptidase.

30. A method of producing a peptide comprising:

cleaving a fusion protein construct, which includes a carbonic anhydrase and a variable fused polypeptide, to produce a soluble carbonic anhydrase fragment and a soluble variable fused polypeptide fragment; and

precipitating the carbonic anhydrase fragment.

31. The method of claim 30 further comprising extracting the precipitated carbonic anhydrase fragment with a solvent to recover an additional amount of the variable fused polypeptide fragment, wherein the solvent includes an organic component.

32. The method of claim 31 wherein the organic solvent solvent comprises acetonitrile, propanol, citric acid, polyethyleneglycol, or mixtures thereof.

33. The method of claim 30 further comprising isolating the fusion protein construct.

34. The method of claim 33 wherein isolating the fusion protein construct includes precipitating the fusion protein construct.

35. The method of claim 30 comprising precipitating the soluble carbonic anhydrase fragment from a solution at a pH of about 3.2 to about 6.0.

36. The method of claim 35 wherein precipitating the soluble carbonic anhydrase fragment further comprises adding a salt to the solution.

37. The method of claim 36 wherein the salt includes a divalent metal cation.

38. The method of claim 33 wherein isolating the fusion protein construct includes contacting a solution of the fusion protein construct with a support which includes a benzenesulfonamide compound or a acetazolamide compound.

39. The method of claim 38 wherein contacting the fusion protein construct solution with the support includes passing the fusion protein construct solution through an affinity column which includes a sulfanilamide compound.

40. The method of claim 33 further comprising expressing the fusion protein construct as part of an inclusion body in a host cell.

41. The method of claim 33 wherein isolating the fusion protein construct comprises adding PEI to a solution of the fusion protein construct.

42. The method of claim 41 further comprising adjusting the pH of the PEI solution to between about 3.2 and

about 6.0, thereby precipitating the fusion protein construct.

43. The method of claim 30 wherein the carbonic anhydrase is linked to the variable fused polypeptide through an enzymatic cleavage site; and

wherein cleaving the fusion protein construct includes treating the fusion protein construct with an endopeptidase selected from the group consisting of enterokinase, Factor Xa, ubiquitin cleaving enzyme, thrombin, trypsin, renin, subtilisin, chymotrypsin, clostripain, and *S. aureus* V8.

44. The method of claim 43 wherein cleaving the fusion protein construct comprises treating the fusion protein construct, which is dissolved in a solution which includes sodium chloride, with thrombin.

45. The method of claim 43 wherein cleaving the fusion protein construct comprises adding a anion exchange resin to a solution of the fusion protein construct and treating the fusion protein construct with enterokinase.

46. The method of claim 30 wherein the carbonic anhydrase is linked to the variable fused polypeptide through a chemical cleavage site; and

wherein cleaving the fusion protein construct includes treating the fusion protein construct with a chemical cleavage agent selected from the group consisting of cyanogen bromide, hydroxylamine, BNPS-skatole, an S-cyanylating agent and an acid having a pK_a of no more than about 3.0.

47. An inclusion body expressed in an *E. coli* host cell comprising a fusion protein construct, wherein the

fusion protein construct includes a carbonic anhydrase and a target peptide, said target peptide including an amino acid sequence corresponding to a peptide selected from the group consisting of GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22).

48. The inclusion body of claim 47 wherein the fusion protein construct includes at least 2 copies of a target peptide.

49. The inclusion body of claim 47 wherein the target peptide includes an amino acid sequence corresponding to PTH(1-34) (SEQ ID NO:22); and wherein the carbonic anhydrase is linked to the target peptide through an interconnecting peptide which includes an enzymatic cleavage site recognized by thrombin.

50. The inclusion body of claim 47 wherein the target peptide includes an amino acid sequence corresponding to GRF(1-41) (SEQ ID NO:23); and wherein the fusion protein construct further comprises an enzymatic cleavage site recognized by an enzyme selected from the group consisting of thrombin, enterokinase and carboxypeptidase Y.

51. The inclusion body of claim 48 wherein the target peptide includes an amino acid sequence corresponding to GRF(1-44) (SEQ ID NO:20).

52. The inclusion body of claim 47 wherein the target peptide includes an amino acid sequence corresponding to GLP1(7-34) (SEQ ID NO:21); and wherein the fusion protein construct further comprises an enzymatic cleavage site recognized by trypsin.

53. The inclusion body of claim 47 wherein the target peptide includes an amino acid sequence corresponding to GLP1(7-34) (SEQ ID NO:21); and wherein the fusion protein construct further comprises a methionine residue.

54. A method of producing a peptide comprising:

expressing a fusion protein construct as a part of an inclusion body in an *E. coli* host cell, wherein the fusion protein construct includes a carbonic anhydrase and a variable fused polypeptide; and

isolating the fusion protein construct.

55. The method of claim 54 wherein the variable fused polypeptide comprises a target peptide which includes an amino acid sequence corresponding to a peptide selected from the group consisting of GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22).

56. The method of claim 54 wherein the variable fused polypeptide includes at least 2 copies of a target peptide.

57. A fusion protein construct comprising a carbonic anhydrase and a variable fused polypeptide, wherein the variable fused polypeptide includes a target peptide, the target peptide including an amino acid sequence corresponding to a peptide selected from the group consisting of GRF(1-41) (SEQ ID NO:23), GLP1(7-34) (SEQ ID NO:21) and PTH(1-34) (SEQ ID NO:22).

58. The fusion protein construct of claim 57 wherein the variable fused polypeptide includes a target peptide corresponding to a peptide selected from

the group consisting of GRF(1-41) (SEQ ID NO:23), GRF(1-44) (SEQ ID NO:20), GLP1(1-34) (SEQ ID NO:26), GLP1(7-34) (SEQ ID NO:21), GLP1(7-36) (SEQ ID NO:27), GLP1(7-37) (SEQ ID NO:28), PTH(1-34) (SEQ ID NO:22), PTH(1-38) (SEQ ID NO:29) and PTH(1-84) (SEQ ID NO:25).

59. A recombinant gene containing a DNA sequence coding for the fusion protein construct of claim 57.

60. An expression cassette comprising:

a nucleic acid sequence coding for the fusion protein construct of claim 57; and

wherein the nucleic acid sequence is operably linked to a promoter functional in a vector.

61. An expression vector comprising the expression cassette of claim 60.

62. A transformed cell comprising a recombinant gene including a DNA sequence coding for the fusion protein construct of claim 57.

63. A method of producing a peptide comprising:

cleaving a fusion protein construct, which includes a carbonic anhydrase and a variable fused polypeptide, to produce a carbonic anhydrase fragment and a variable fused polypeptide fragment;

precipitating the carbonic anhydrase fragment and the variable fused polypeptide fragment; and

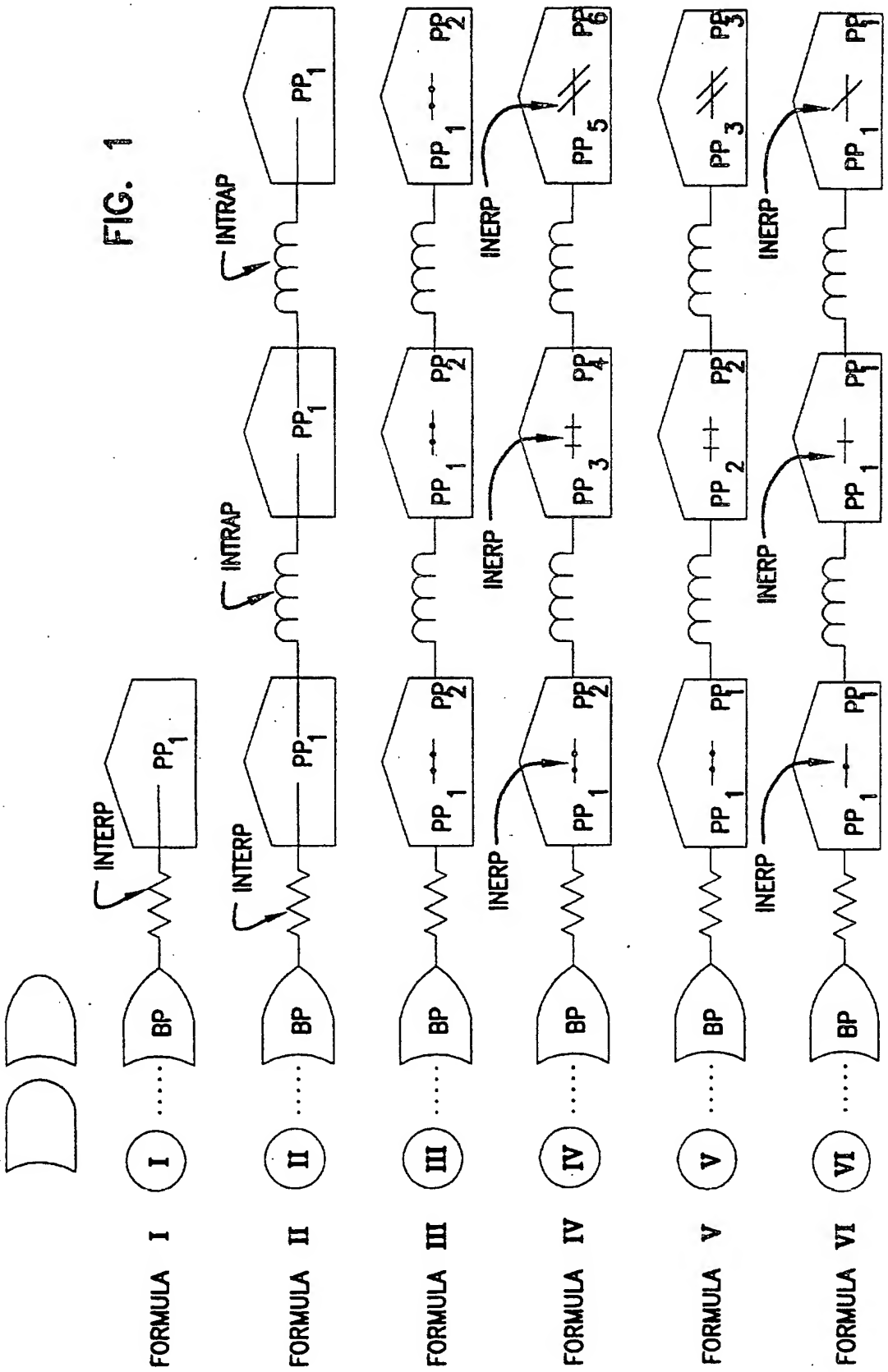
extracting the precipitated fragments with a solvent which includes an organic component, thereby recovering the variable fused polypeptide fragment.

64. The method of claim 63 wherein the organic solvent solvent comprises acetonitrile, propanol, citric

acid, polyethyleneglycol, or mixtures thereof.

65. The method of claim 63 wherein the variable fused polypeptide includes at least two copies of a target peptide.

FIG. 1


FORMULA I
FORMULA II
FORMULA III
FORMULA IV
FORMULA V
FORMULA VI
BP
INTERP
INTRAP
INERP
PP1
PP2
PP3
PP4
PP5
PP6

INTERNATIONAL SEARCH REPORT

International application No
PCT/US 95/15800

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/62 C12N9/88 C07K14/60 C07K14/605 C07K14/635
C07K1/30 C07K1/113

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO,A,92 01707 (BIONEBRASKA, INC.) 6 February 1992 cited in the application see page 3, line 20 - page 7, line 2 see page 7, line 14 - page 11, line 22 see page 16, line 25 - page 17, line 18 see page 20, line 2 - line 18 see page 22, line 24 - page 25, line 3 see page 32, line 30 - page 40, line 4 see page 44, line 13 - line 36 see page 51, line 25 - page 52, line 6 see page 55, line 8 - page 58, line 32; examples 6-8 --- -/--	47-62

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
22 April 1996	10. 05. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Montero Lopez, B

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP,A,0 528 686 (SUNTORY LTD) 24 February 1993 see page 3, line 13 - line 28 see page 4, line 1 - line 6 see page 9, line 48 - page 11, line 35 see page 13, line 44 - page 14, line 27	1-65
A	WO,A,88 07086 (CREATIVE BIOMOLECULES, INC.) 22 September 1988 see page 4, paragraph 1 - page 5, paragraph 1 see page 24, paragraph 3 - page 25, paragraph 1 see page 30, last paragraph - page 32, paragraph 1	1-65
A	WO,A,92 00993 (CALIFORNIA BIOTECHNOLOGY INC.) 23 January 1992 see page 6, line 9 - page 7, line 9 see page 13, line 10 - page 15, line 24	1-65
A	PROTEIN EXPRESSION PURIF. (1993), 4(4), 265-74 CODEN: PEXPEJ;ISSN: 1046-5928, 1993, XP000568594 VAN HEEKE, GINO ET AL: "Gene fusions with human carbonic anhydrase II for efficient expression and rapid single-step recovery of recombinant proteins: Expression of the Escherichia coli F1-ATPase.epsilon. subunit" see page 266, left-hand column, paragraph 2 see page 267, left-hand column, paragraph 2 - page 268, left-hand column, paragraph 1 see page 272, right-hand column, paragraph 4 - page 273, right-hand column, paragraph 3	1-65
A	WO,A,88 07085 (CREATIVE BIOMOLECULES, INC.) 22 September 1988 see page 6, paragraph 1 - paragraph 2 see page 10, paragraph 2 - page 11, paragraph 2 see page 27, paragraph 2 - page 29, paragraph 2 see page 30, paragraph 2 - page 31, paragraph 1 see page 36, paragraph 2 - page 37, paragraph 1 see page 39, last paragraph - page 40, paragraph 2 see page 45, last paragraph - page 46, paragraph 2	1-65

1

INTERNATIONAL SEARCH REPORT

Int. ation on patent family members

Internatior Application No
PCT/US 95/15800

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9201707	06-02-92	AU-B-	662302	31-08-95
		AU-B-	8915591	18-02-92
		EP-A-	0539530	05-05-93
		JP-T-	6505624	30-06-94
EP-A-0528686	24-02-93	AU-B-	656791	16-02-95
		AU-B-	2107892	11-03-93
		CA-A-	2076320	20-02-93
WO-A-8807086	22-09-88	AU-B-	601273	06-09-90
		AU-B-	1482488	10-10-88
		CA-A-	1334943	28-03-95
		DE-D-	3884529	04-11-93
		DE-T-	3884529	28-04-94
		EP-A,B	0305481	08-03-89
		JP-T-	1503117	26-10-89
		US-A-	5330902	19-07-94
		US-A-	5302526	12-04-94
		US-A-	5215896	01-06-93
WO-A-9200993	23-01-92	US-A-	5258496	02-11-93
		CA-A-	2065323	11-01-92
		EP-A-	0491033	24-06-92
		JP-T-	5502888	20-05-93
		US-A-	5403915	04-04-95
WO-A-8807085	22-09-88	AT-T-	113985	15-11-94
		AU-B-	605291	10-01-91
		AU-B-	1575088	10-10-88
		DE-D-	3852074	15-12-94
		EP-A-	0305500	08-03-89
		EP-A-	0621340	26-10-94
		JP-T-	2500876	29-03-90
		US-A-	5013653	07-05-91